April 29, 2010

What to do about TBTF?

Mark J. Flannery*
University of Florida
Warrington College of Business
flannery@ufl.edu

ABSTRACT

Government actions during the recent financial crisis have ratified the prior expectation that sufficiently large or complex financial institutions are too big to fail. The government’s guarantee will permit TBTF institutions to borrow more cheaply, to earn more profits, and to expand at the expense of un-guaranteed institutions. Competition among TBTF lenders will probably transmit risk-taking subsidies to their borrowers and thence into the real sector. The root cause of TBTF government interventions is the disruption that would accompany a large institution’s bankruptcy filing.

Proposed reforms treat TBTF-related risk-taking as an externality. Economists’ generally prefer to deal with externalities through pricing, which in this instance could be implemented as risk-based capital requirements or insurance premia. However, measurement errors may make it appropriate to treat these externalities in part with quantity constraints (“Volcker” restrictions). I propose that we limit TBTF distortions via a combination of direct risk-restrictions (inside the firm and in the financial infrastructure) and an aggressive set of risk-based capital requirements.

No matter how we decide to curtail TBTF externalities, we must recognize that some of the excluded risks will migrate to other, less regulated financial entities. This complicates evaluating the impact of TBTF risk reduction on financial stability.

This paper was written while I was Resident Scholar at the Federal Reserve Bank of New York. The opinions expressed are entirely my own, and reflect no official positions of the New York Fed or the Federal Reserve System. I thank Ron Feldman, Beverly Hirtle, Alan Blinder, and Andy Haldane for suggestions and comments on the general topic of this paper. I am also indebted to Larry Wall for his insightful comments on the first draft of this paper. Remaining errors are my own.
“[O]ne of the greatest threats to the diversity and efficiency of our financial system is the pernicious problem of financial institutions that are deemed ‘too big to fail.’ … It is unconscionable that the fate of the world economy should be so closely tied to the fortunes of a relatively small number of giant financial firms.” (Bernanke (2010), pages 1, 3)

The idea that some financial firms are TBTF dates back at least to the 1984 rescue of Continental Illinois. The Comptroller subsequently told Congress that financial supervisors would not permit any of the eleven largest U.S. banks to fail (Carrington (1984)). O’Hara and Shaw (1990) conclude that the Comptroller’s statement raised the largest U.S. banks’ stock prices, with the gain being largest for the weakest banks. In other words, the government’s guarantee provided tangible value to bank shareholders. The distortions associated with deposit insurance were already well-known (e.g., Merton (1977)), but the Comptroller’s announcement extended the range of these potential distortions. Recent government support for distressed financial institutions has further broadened the safety net.

The financial and real-sector distortions generated by guaranteeing bank deposits or liabilities require offsetting government regulation (Flannery (1982)). In the 1980s, deposit insurance premia were the same for all banks, so this regulation took the form of capital adequacy requirements and preventing banks from undertaking certain sorts of risky business. The Bank Holding Company Act was administered under the belief that the banking entity was the main thing to be protected. What holding companies did via other subsidiaries was not so important as what the bank did – so long as the bank could be properly insulated from the affairs of the other subsidiaries.

The recent financial crisis illustrated the inadequacy of this regulatory view, and of the manner in which regulations were implemented. Size or complexity elicited government support, largely independent of the troubled firm’s institutional form. Government financial support was extended to large bank holding companies, investment banks, and insurance companies via equity purchases and broad credit guarantees. Money market mutual funds were guaranteed, the Fed purchased commercial paper, Congress raised the FDIC insurance limit, and the Fed began accepting a wider range of collateral
for discount window loans to a broader range of institutions. Even the government’s decision to withhold support from Lehman Brothers illustrated how much the markets had come to depend on government guarantees: when Lehman failed, short-term credit markets collapsed and financial firms’ share values fell precipitously.

I do not wish to second-guess official decisions made in the face of unprecedented financial turmoil. These actions averted a worse financial collapse. However, those decisions starkly illustrate the time consistency problem underlying financial policy: an imminent disaster was mitigated, but the rescue has serious implications for the form of our financial system going forward. The widespread conjecture that large financial firms would be protected from failure was confirmed in 2008. Subsequent policy discussions of how “systemically important financial institutions” (SIFIs) should be treated to prevent a repeat of the 2008 debacle add credibility to the notion that most of the liabilities issued by major financial institutions enjoy a full de facto guarantee. The question that arises along with these government guarantees is how to reform the guaranteed institutions’ operational characteristics to prevent a repeat of the recent bailouts, socialized private losses, and dysfunctional lending behavior.

This paper evaluates the policy alternatives available for minimizing the financial and real-sector distortions resulting from TBTF policies, whether real or conjectured.¹ Section I identifies the source of the problem – the disruptions of a threatened traditional bankruptcy for firms whose credit worthiness is crucial to their business operations. Section II describes the problems caused by a TBTF policy. Section III considers potential policy responses from a broad theoretical perspective. Section IV interprets some of the proposed solutions -- including size limits, activity limits, new fees on large firms, and revisions to trading and bankruptcy procedures -- in the context of the theoretical properties discussed in Section III. Section IV recommends the most appropriate policy responses to the TBTF problem, and the final section concludes.

¹ I had intended to consider the GSEs in this paper, but found that the associated issues relating to U.S. housing policy could not be accommodated in the allotted space. Still, their resolution and re-formation remain important issues that contribute to the challenges of managing TBTF guarantees going forward.
I. The Source of "TBTF"

Merton (1990) distinguishes between a firm’s “customers” and its “investors”. Customers wish to purchase a firm’s products, while investors provide a firm with the financial resources to produce those products. Unlike customers, investors should hold well-diversified portfolios of claims, so a poor realized return from any one investment should not threaten their survival. The typical non-financial firm has separate sets of investors and customers. So it can continue serving customers while a lengthy bankruptcy proceeding re-arranges investors’ claims. Moreover, most non-financial firms do not use a substantial amount of short-term liabilities. A troubled non-financial firm therefore generally has time to seek a solution to its financial difficulties, either in bankruptcy or through some type of exchange offer.

For many financial products, customers must become creditors of the firm providing those products.

- Banks generally require that customers make payments from demand deposit accounts with positive balances.
- Over-the-counter (OTC) counterparties in foreign exchange and in swaps must accept credit risk exposures to one another.
- Lenders involved in repurchase (RP) agreements want the liquidity promised by their contract, and not the securities underlying that promise.
- Customers seeking credit guarantees or lines of credit rely on their banks’ future solvency.

A bank’s financial distress leads customers to move their business elsewhere, which exacerbates the troubled bank’s need for liquidity. Because many of the relevant claims are short-term, such changes can happen rapidly. A possible bankruptcy is particularly frightening to bank customers because it raises questions about access to collateral, rights of offset, netting in derivatives positions, etc. These questions loom even larger when the bank’s assets are scattered around the world, in a variety of political jurisdictions. At a minimum, bankruptcies take time, and the loss of liquidity during that time would be inconvenient or expensive. In short, the threat of a financial firm’s insolvency causes rational customers

---

2 For example, see Wallison (2010, pages 11-12) or Baxter and Sommers (2010).
to flee, destroying the firm’s charter value and forcing some immediate resolution to the solvency questions.

In the recent financial crisis, policymakers were unwilling to impose losses even on financial firms’ junior-most creditors because such impositions required an extraordinarily disruptive bankruptcy proceeding. Worldwide, hybrid debt instruments (including U.S.-type subordinated debentures) provided no loss-absorption capacity. Going forward, official unwillingness to permit hybrid debt to bear losses also renders debt market discipline questionable at best.

The largest, most inter-connected financial firms received particular protection because policymakers feared substantial “contagion” or “knock on” effects from their failure. If a large bank’s losses would cripple enough other financial firms, nonfinancial firms would lose access to the funding they require to continue operating. The mechanics of such contagion could come from several sources. First, large banking firms would be likely to have credit relationships with a large number of counterparties. If the counterparties were risk-averse, they would diversify their exposures to protect against the default of any one. However, conjectured guarantees of the largest banks would make counterparties willing to accept large exposures to such firms. A belief in TBTF institutions can thus neutralize normal tendencies toward market discipline and make large firms so “interconnected” that their default would harm many institutions and investors.

A second cause of undiversified counterparty exposures could arise when a large bank provides a unique service that requires customers to take on credit risk. Only a few banks operate active derivatives trading desks or make markets in certain types of securities. A few more undertake substantial FX trading.3 Because active, centralized markets have economic benefits (Pagano (1989)), we should not blithely decide to break up those activities as a means of reducing SIFIs’ systemic importance.4

---

3 In principle, most such credit risks could be collateralized. Market participants have been working to limit the impact of counterparty risk on OTC transactions, without (so far) eliminating it entirely. Presumably, further counterparty risk mitigation is not privately optimal.

4 Another example of concerned customer exposures involves hedge funds, which tend to concentrate their dealer services with one prime broker in order to minimize their net collateral requirements.
Finally, large SIFIs are probably more diversified than their smaller counterparts. So if a single risk threatens the solvency of a SIFI, there may be other firms with equal (or greater) exposures to the same risk. Forcing the SIFI into failure can generate a fire sale that imposes losses on other firms with similar assets (Brunnemeier and Pedersen (2009)). The SIFI’s failure might also induce runs at other institutions.5

When the threat of insolvency can cause a financial firm’s demise, it may be good policy to insulate private creditors from loss during times of financial turmoil. Going forward, it seems that the government actions of 2008 have left a legacy of strongly conjectured guarantees for SIFIs, probably even in the absence of a general turmoil.

II. Consequences of TBTF

The effects of subsidized risk-bearing are well known from numerous discussions of federal deposit insurance. Guaranteed institutions enjoy lower borrowing rates because investors do not require compensation for possible default losses. (At least, investors require a lower default risk premium than they would in the absence of government guarantees.) This borrowing subsidy occurs whether the guarantees are de jure or merely conjectured.6

Equivalently, government insurance makes a financial institution’s borrowing cost less sensitive to leverage or asset volatility increases.7 SIFIs therefore encounter far less market pressure to raise capital cushions, and have a comparative advantage in competing for higher-risk assets. While government guarantees increase TBTF firms’ default probabilities, they also make it possible for TBTF

---

5 Aharoney and Swary (1996) identify five large bank holding companies (BHC) that failed in the southwestern United States between 1986 and 1989. They conclude that these failures lead investors to revise their valuations of other banks thought to have similar portfolios, but had no significant effect on banks with other risk exposures.

6 Some empirical evidence indicates that bank value increases when it becomes large enough to be considered TBTF (Brewer and Jagtiani (2010), Penas and Unal (2004)).

7 Merton (1977) shows that deposit insurance unambiguously increases an insured bank’s desired portfolio risk and leverage in a one-period model. Marcus (1984) shows that a positive charter value will limit the firm’s interest in exploiting mispriced deposit insurance, but not eliminate it entirely. Theory cannot determine unambiguously whether the increased default probability results from higher leverage or higher asset. The important point is that SIFIs will choose a higher default probability on account of deposit insurance.
institutions to expand at the expense of their smaller competitors, presumably mis-allocating investable savings across geographic regions. Competition among SIFIs tends to pass through some of the risk-taking subsidy to their customers, thereby distorting real sector activities. To the extent that insurance alters SIFIs’ preferred portfolio risks, they may all seek loans (activities) with similar risk characteristics. If SIFIs accumulate similar risk exposures, the probability of another systemic crisis rises, with large potential costs for taxpayers.

Market discipline from bank creditors is another victim of a TBTF. If large liability-holders felt their funds were at risk, they would monitor banks more carefully and demand an appropriate premium for bearing risk. Although these firms are complex and can change their risk exposures rapidly, considerable evidence supports the hypothesis that bank claimants can be effective monitors (Flannery (2010)). Without TBTF guarantees, financial supervisors could rely on information generated by monitoring creditors to improve their assessments of regulated firms’ condition.

Unfairly (incompletely) priced liability guarantees lead insured firms to generate negative financial externalities by operating with inefficiently high default probabilities. Because large financial firms cannot be put through a standard bankruptcy process, supervisors tend to absorb a distressed firm’s losses on behalf of its liability-holders. “Systemically important” firms, by definition, impose further external costs if they fail, on parties with which they may have no direct relationship.

Standard ways to offset an externality include pricing (taxing) it and restricting its quantity. “Pricing” implies a risk-related insurance fee (in the manner of FDIC insurance premia) or a risk-based capital standard.8 Either tool can be used to offset TBTF competitive benefits. A risk-based insurance premium that mimics the credit risk premium investors would charge in the absence of a government guarantee induces non-SIFIs to employ the socially optimal amount of leverage. Alternatively, the government insurer can charge a fixed premium and constrain leverage (in a risk-sensitive manner) to drive the TBTF firm’s liability cost to include a fair risk premium. However, quantity restrictions can

8 The Obama administration would levy a fee on banks’ uninsured liabilities (except deposits) as a means of discouraging size and recouping TARP expenditures. Perotti and Suarez (2010) would levy a fee on short-term liabilities to discourage leverage.
also control externalities. This is the logic behind the so-called Volcker rule, which would prevent a SIFI from expanding into areas that are considered particularly high-volatility (perhaps investment banking or proprietary trading).

“Systemic” external effects of SIFI failures would require that SIFIs be made less risky than their smaller counterparts. This can be accomplished by imposing insurance premia or capital requirements that exceed the levels that produce “fair” (market equivalent) pricing. SIFIs’ default probabilities can also be constrained by more restrictive risk-taking limits.

III. Policy Responses: Theory

If supervisors could observe the firm’s selected asset volatility and its market-valued leverage, pricing can constrain a firm’s default probability to a socially optimal level. Likewise, if regulators can observe and control an insured institution’s ability to take on risks, “quantity” restrictions on bank risk-taking could produce a first-best equilibrium. Figure 1 illustrates the equivalence of pricing and quantity solutions to the problem of correcting insured institutions’ externalities.

Most economists prefer to ration scarce commodities via price, and instinctively choose risk-based premia or capital requirements as the most appropriate way to control SIFI risk-taking. Pricing the externality permits individual firms to reduce risks in a cost-efficient manner. In addition, risk-based insurance fees can generate a fund to handle default losses.\(^9\) However, policymakers have only imperfect information; they cannot perfectly determine the amount of risk associated with any fee level. In Figure 2, supervisors do not know whether the SIFI sector’s response to risk-based pricing is described by the schedule “A” or “B”. A fee (\(F^*\)) therefore reduces the distortions caused by TBTF policies, but the fee does not yield a known quantity of risk: for \(F^*\), the firm’s default risk could be \(Q_A\) or \(Q_B\). Too low a fee leaves SIFIs subsidized and excessively risky; too high a fee drives risk outside the regulated sector. In contrast, imposing a quantity restriction (\(Q^*\)) would yield just the desired amount of risk if policymakers’

\(^9\) Some observers consider such a fund useful. Others worry that it will be used to subsidize all failures, regardless of the individual situation.
uncertainty took the form indicated in Figure 2. However, supervisors might have only a noisy assessment of how quantitative restrictions will affect actual SIFI risk, as indicated by the range of measurement error on the horizontal axis in Figure 3. Quantity restrictions thus might also produce an uncertain equilibrium risk level.¹⁰

Once we allow for inevitable measurement errors, pricing or quantity instruments both yield imperfect control. The optimal tool(s) will depend on the sources and magnitudes of policymakers’ uncertainty.¹¹ At this point I don’t wish to argue for either pricing or quantity controls, but only to illustrate how unavoidable measurement errors might make the usual preference for pricing externalities inferior.

Note further that both control mechanisms seek to drive some risks out of the SIFI system. Will this necessarily increase financial stability? Litan (1987), Pierce (1991) and others once suggested that deposit insurance be provided only to banks that held risk-free assets. They argued that this would make deposit insurance effectively free to provide. Depositors could get safe investments and safe access to the payments system, without government subsidies. A powerful argument against narrow banks recurs in the context of SIFI risk-taking: many of the risky activities forced out of insured banking institutions (including maturity mismatching and high-risk lending) would migrate elsewhere. (We didn’t know at the time to call these new firms “shadow banks.”) If supervisors would let shadow banks fail, narrow banking would provide a sound solution to distorted risk-taking incentives. But if the financial services provided by nonbanks were deemed systemically important (as they have been during the recent crisis), they would operate with a conjectural guarantee. The distortive effects of government protection would be associated with different institutions, but would not be eliminated from the system.

¹⁰ Monetary economists will recognize the similarity of my discussion here to Poole (1970).
¹¹ Over time, supervisors should learn to assess the impact of their fees and/or quantity restrictions. However, if the banks are also changing over time there is no reason why measurement errors should shrink over time, much less disappear.
Likewise today, we need to consider what will happen to the risks squeezed out of the banking system. Will their destinations feed back to affect financial stability in dysfunctional ways? I don’t have any further insights into this issue, except to emphasize that “out of sight” is not the same as “gone”.

IV. Policy Responses re: Individual Firms

The potential for measurement errors in assessing risk exposures suggests that we should not expect a single policy change to set things right. With measurement errors, some combination policy is generally the best way to mitigate externalities. This section considers the various suggested mechanisms for making SIFIs less failure-prone. The following section discusses changes to the institutional environment, and the paper concludes by offering a recommended approach to reducing TBTF external effects.

A. Better Supervision

Supervisors may have had most of the regulatory tools they needed to prevent the worst of the recent financial crisis. However, they failed to act forcefully. The important question is why supervisors did not act sooner, for example to curtail purchases of subprime loans or to require prompt recapitalizations when losses began to mount. Perhaps the fault lay with particular individuals, or with a faulty oversight that can be improved. Or perhaps slow actions are unavoidable as a practical matter and we should not expect much help from better future supervision. Certainly some knowledgeable commentators incline toward the latter possibility:

“The belief that appropriate regulation can ensure that speculative activities do not result in failures is a delusion.” Mervyn King (October 21, 2009, page 7)  

The country should "not be overoptimistic about how successful improved oversight will be." (Mankiw (2010))

Taking strong supervisory actions against specific firms is very difficult. Amidst great uncertainty, supervisors must gather enough evidence to support a confident assessment that a firm’s condition is weak. This assessment must be sufficiently strong to resist the firm’s counter-arguments, which will often take the form of “let’s wait and see how things develop.” As illustrated in Figure 3, supervisors
probably have only an imprecise assessment of an institution’s true risk exposure. It is therefore unreasonable to expect that they will take strong, early actions against a specific firm on the basis of such an assessment.

The recent crisis provides a typical example. Even before 2007, some forecasters had predicted that subprime lending would come to grief. However, an even larger number of opinions held that the “new” situation was “unprecedented” and would work as intended. Anecdotal evidence strongly suggests that many financial firms’ risk managers were replaced when they did not understand the Brave New World. Similarly, supervisors only slowly concluded that the apparent dangers of subprime paper were real enough to warrant specific actions. Even the purest, most diligent public servant will be relatively slow to act when events are confusing.

Supervisors’ difficulty identifying true risk exposures may explain part of their support for compensation reform. The notion that compensation arrangements encourage systemically dangerous risk-taking implicitly assumes that supervisors cannot recognize risk accurately in real time. Otherwise, supervisors could implement prompt corrective measures and executive compensation could be left to corporate governance, rather than being a high-profile supervisory problem.

Supervisory controls are further hampered by a regulatory system that defines bank condition largely in terms of a book capital ratio. By construction, book values lag changes in a firm’s valuation. As a firm’s condition deteriorates, market values fall with respect to book values because managers can exercise GAAP options to report uncertainties in a relatively favorable light. Kuritzkes and Scott (2009) point out that

The five largest US financial institutions subject to Basel capital rules that either failed or were forced into government-assisted mergers in 2008 – Bear Stearns, Washington Mutual, Lehman Brothers, Wachovia and Merrill Lynch – had regulatory capital ratios ranging from 12.3 per cent to 16.1 per cent as of their last quarterly disclosures before they were effectively shut down.

---

12 For example, supervisors felt unable to force banks to suspend their dividend payments even at the depth of the crisis when equity market values were historically low. The Basel Committee has proposed rules for “capital conservation” that will expand supervisory powers to avoid such circumstances in the future (Basel Commission on Bank Supervision (2009a), page 9).
Figure 4 plots average equity ratios and CDS spreads for all of the SCAP institutions (except GMAC, which is not publicly traded), between 2004 and early 2009. Note the book equity value’s stable relationship to total assets through this period. Book equity ratios show scant change even as equity’s market value ratio and CDS spreads begin to deteriorate after mid-2007.

The point here is not that markets never make mistakes. Mis-valuations clearly occur, just as some supervisory decisions turn out to be wrong. However, market capital measures are nearly always more accurate when a bank’s leverage has risen sharply, and supervisors should therefore be free to act on market valuations when they seem more appropriate. The regulations’ reliance on book capital adequacy measures made it difficult for supervisors to mandate dividend restrictions, even as the firms’ market equity ratios were plunging. In addition to the failed institutions cited by Kuritzkes and Scott, several other large institutions suffered dangerous liability runs even as their book capital ratios remained well above “adequate” levels.

The effects of relying on book solvency measures are further illustrated in Figure 5, which plots the Tier 1 ratio and market equity ratios (market value of equity over book value of assets) for individual banks. Citigroup’s Tier 1 ratio held steady from early 2004, despite a nearly monotonic decline in its market equity ratio. After the crisis began, Citigroup issued some new capital instruments, raising its Tier 1 ratio quite markedly, while equity’s market value fell even more markedly. The case for BankAmerica is slightly different, in that its market equity ratio remained relatively high until 2007Q3. From that point on, BankAmerica’s equity value changes tell quite a different story from its Tier I ratio changes. The disparity between market and book ratios is even more dramatic for Wachovia, which was acquired by Wells Fargo in 2008Q3. Wachovia’s market decline began in 2006Q4, following its $25.5 billion acquisition of Golden West Financial in mid-2006, and continued unabated.

The final panel in Figure 5 illustrates one of the crisis’ strong survivors. JPMorgan/Chase’s market value began to fall in early 2007, but its decline was smaller than it was for the other firms’ plotted here: from 12.39% in 2006Q4 to a low of 5.02% two years later. By contrast, Citigroup’s and

---

13 Thanks to Kevin Stiroh for permitting me to use his graph.
BankAmerica’s market capital ratios fell more than 12% over the same two-year period. These examples clearly illustrate how book solvency measures provide a distorted assessment of firm condition when the firm is under stress.

In summary, we should heed Messrs. King’s and Mankiw’s admonitions about the limits of supervisory oversight. Supervisory processes can probably be improved, but the hope that a SIFI’s problems can generally be headed off by discretionary supervision seems unwarranted. Better supervision will not save us from future crises. The problem is just too hard.

B. Size Limits

Numerous writers have suggested the ultimate “quantity” restriction for TBTF firms: that they are too big to exist, and should be broken up (Reich (2008, 2010), Shultz (2008), O’Driscoll (2009), Johnson (2009)). The underlying motivations include both economic and political considerations.

Economically, the strongest case for limiting an institution’s connectedness probably involves the services sold to “customers” who absorb a relatively undiversified credit risk exposure along with a service like trading and market-making. At the same time, trading exhibits scale economies: a market becomes more liquid when it attracts more traders, reducing the cost of transacting (Pagano (1989)). Policymakers thus confront a tradeoff between efficiency in the provision of trading services and interconnectedness. Furthermore, if U.S. providers of trading services contract, the scale economies may move abroad, placing U.S. banks at an international comparative disadvantage.14

Larger banks might have greater systemic effects simply because they deal with more customers, across a broader geographic area. So a very large bank’s failure affects many parts of the economy and bad events might feed upon one another. Furthermore, a large bank may more successfully engineer a government rescue because it has a larger base of endangered voter-customers. (That is, size may correlate with the extent of conjectured guarantees.) Either of these situations can be addressed via a fee levied on large banks’ liabilities or a higher capital standard. It’s not clear why breaking up large firms – trying to set “Q” directly by policy – should be more appropriate.

---

14 Even if this situation were economically desirable, political forces would certainly act to protect the status quo.
Even if one were inclined to address size-related issues with a quantity constraint, a number of important questions remain about implementation (Stern and Feldman (2009)). Would large (or interconnected) firms be required to shrink across-the-board, or could they choose which activities to curtail? Their choices would doubtless reflect scale economies, so trading operations might not be the first activities spun off. What if the shrinkage did not address the issues relevant to policymakers?

A policy of simply dismantling large firms seems too radical and imprecise when the social costs of large institutions can be addressed via pricing,

\[ \text{C. “Volcker” Restrictions} \]

Paul Volcker’s name has been attached to a variety of proposals that would limit SIFIs’ permissible activities.\(^{15}\) The “Volcker” remedies thus focus on quantity restrictions. The broad purpose is to keep SIFIs out of the most volatile (socially dangerous) activities. The logic in this position reflects a change in bank portfolio risks over time. When Basel I was first implemented, U.S. banks were basically restricted to hold fixed-income claims, financed largely with short-term deposits. Subsequent statutory revisions permitted banks to enter the underwriting, asset management, and trading businesses, where returns are more equity-like. (A big step in this direction – but not the only one – was the Gramm-Leach-Bliley Act in 1999, which “repealed Glass Steagall”.) When an investment offers large upside potential gains, an investor is more willing to accept large potential losses, which increase the risk of bankruptcy. Furthermore, combining high-variance returns with a more traditional banking book exacerbates the supervisor’s problem in identifying true risk exposures, especially when those risks can change rapidly in the normal course of business.

The most obvious instance of this substitution is proprietary trading, including banks’ participations in hedge fund investments. Appropriate incentives for traders in such accounts are very difficult to implement, generating a bias toward higher-risk positions. Combined with banks’ higher leverage and short-term funding, high-volatility investments generate the potential for frequent distress.

\(^{15}\) Volcker’s suggestion was first presented in Group of Thirty (2009). Other senior financial sector participants quickly expressed support for limiting bank activities (Uchitelle (2010)).
Preventing banks from operating proprietary trading desks thus seems like a reasonable “Q” type restriction. In principal, liability fees or capital requirements could be calibrated to offset a SIFI’s distorted risk-taking incentives. But in fact, proprietary trading risks can change so rapidly and unobservably that price-based restrictions seem doomed to fail in this context.

At the same time, identifying “proprietary” trading is fraught with difficulty. Bank market-making activities require liquidity provision, which in turn requires banks to hold some securities when they don’t have an immediately offsetting market order. Strictly limiting bank trading “for its own account” thus interferes with the efficiency of OTC trading. Yet permitting banks to provide market liquidity leaves them holding investments that could have highly volatile returns. A balance must be attained in restricting proprietary positions, unless SIFIs are to be excluded entirely from valuable market-making activities.

Some commentators have proposed that something like the Glass-Steagall separation of investment from commercial banking should be re-imposed, reversing the GLB liberalization of 1999. However, few discussions specify exactly which banking components should be eliminated under a revised Glass-Steagall. This is important because “investment banking” activities entail varying risks. Underwriting, for example, was limited before GLB in 1999, yet this is a relatively sedate business and most consistent with banks’ financing services to nonfinancial corporations. Trading and market-making are less sedate, as discussed above. But it is not clear that “repealing Glass Steagall” would eliminate these activities from SIFIs. Banks have long operated trading accounts, through which they purchase a variety of securities and derivatives, making markets in some. It seems likely that many of the riskier trading activities can be undertaken through a bank trading account as well as through a securities subsidiary. This makes the “investment banking” ability to hold and trade securities seem less important as a unique source of risk. Finally, asset management services constitute part of investment banking, although one might expect banks to offer related services through trust departments (or trust subsidiaries).

In the current crisis, a large proportion of total losses occurred in the large banks’ trading books (BCBS (29009b, page 1)). Trading account capital requirements concentrated on market risk factors
(99%/10-day VARs), ignoring most of the problems that emerged with holding “toxic assets”: the potential for default, credit migrations, illiquidity, and widening spreads. The capital standards for market risk were therefore much lower than the 8% required capital for loans held in the banking book. The ability to hold similar securities in either a banking or a trading book confounded supervisory efforts to measure risks and impose appropriate capital standards. This may have been the greatest social cost of permitting trading activities to co-exist with a traditional banking book: managers thereby acquired a golden opportunity for arbitraging capital requirements. The Basel Committee has proposed a new method for computing required capital against trading positions, emphasizing the relevance of credit and migration risks (Basel Committee on Bank Supervision (2009a)). But the prohibition against proprietary trading and investment banking must confront activities that could still be housed in a bank’s trading account.

V. Improving Systemic Processes

The externalities associated with a TBTF policy can be addressed in the financial system’s infrastructure, as well as within each SIFI individually. Think of firm-specific policies as addressing primarily the SIFIs’ “probability of default” (PD), while infrastructure reform aims to reduce the “loss given default” (LGD). Two areas of potential reform have been offered: enhanced resolution authority for failing SIFIs and moving OTC settlement from a bilateral process to a centralized counterparty.16

A. Resolution Revisions

The Obama administration has requested enhanced resolution powers for both bank and non-bank SIFIs. The goal is to make the reorganization process for over-levered financial firms quicker and more predictable than a traditional bankruptcy filing would be. The maintenance of “living wills” might facilitate such a rapid reorganization outside of normal bankruptcy channels. If supervisors could quickly de-leverage a large financial firm by imposing losses on junior debt and replacing some debt with equity

---

16 A central counterparty only records and settles transactions, which are the most important actions for controlling systemic risk. An “exchange” also publicizes transaction prices, which threatens dealer profits.
(e.g. through exchanges), short-term claimants would have less incentive to run as a firm’s condition weakened. In order to avoid the flight of short-term lenders, however, the set of loss-bearing debt must be readily identified, and its amount must be sufficient to permit a meaningful increase in the firm’s capital ratio. Short-term claimants must be protected from haircuts, in order to preserve the firm’s charter value. Supervisory discretion about which claims get haircuts would be detrimental to this process, which should be rapid and as transparent (predictable) as possible.

In principle, a non-bankruptcy reorganization channel for SIFIs makes a lot of sense. But the complexity of SIFIs’ organizational structures introduces some serious problems. Not only do SIFIs operate with a bewildering array of subsidiaries (Cummings and Eisenbeis (2010)), but they generally operate in many countries. Without very close coordination of resolution decisions across jurisdictions, a U.S. government reorganization would likely set off a scramble for assets of the sort that bankruptcy is meant to avoid. Rapid asset sales could generate downward price spirals (Brunnermeier and Pedersen (2009)), with systemically detrimental effects. Second, supervisors would have to assure that SIFIs maintain the proper sort and quantity of haircut-able liabilities outstanding. Once a firm has been identified as systemically important, this may be a relatively straightforward requirement to impose, but there remains the danger that “shadow” institutions will become systemically important, before they are properly regulated. (This is not a danger unique to the question of resolution.)

I conclude that the international coordination required to make prompt resolution feasible for SIFIs is a long way off, if it can be achieved at all. Cohen and Goldstein (2009) likewise view international cooperation as crucial:

The Lehman bankruptcy demonstrated that a strengthened resolution process in any single country will have only limited efficacy if there is no strengthened international resolution process. For this reason, it is essential that the Basle Committee on Bank Supervision expedite its work on

---

17 For example, Wallison (2010) says that the prevalence of globally active financial firms and the speed with which financial assets can be moved across national boundaries suggests that major reforms are necessary to reduce what are likely to be highly arbitrary outcomes in any resulting insolvency under a territorial system. And, as noted above, it is virtually impossible in a territorial system for a firm to predict how a particular country will treat its assets in the event of insolvency, or indeed whether there will be any assets in that jurisdiction when insolvency occurs. (page 2).
developing recommendations for cross-border resolution. Although this is our last point in the chronology of this paper, it has the very highest priority and importance. (page 18)

Given the present state of affairs, a forced re-capitalization outside of bankruptcy seems unlikely to work very well for a complex, international firm. The hope offered by this approach is therefore limited to a small number of SIFIs with primarily domestic operations.

B. Exchange-Settled OTC Contracts

The fragility of AIG’s Financial Products Group motivated extraordinary Federal Reserve support primarily because supervisors could not determine which other firms might be harmed if AIG could not honor its CDS obligations. Across the range of derivatives financial institutions trade over the counter, worldwide, the potential credit risk exposure is immense. Bilateral OTC contracts provide valuable economic transformations accompanied by the risk that one’s counterparty will fail to perform. And reversing an OTC contract’s economic effects often requires an investor to take an offsetting position with a third party, which doubles his credit risk exposure. Each party takes costly steps to mitigate counterparty risk.

Market participants have innovated margin-posting and marking-to-market processes over the years, to mitigate counterparty credit risks. However, considerable counterparty risk remains in the system, leading many observers to propose that OTC derivatives be settled on a central exchange, such as the CME. (A similar situation existed with respect to “Herstatt risk” in FX settlements, which was eventually addressed (with regulatory encouragement) by formation of the CLS Bank.) Standard contracts could easily be subjected to a margining regime, but highly specialized and idiosyncratic contracts would remain OTC.\footnote{The Squam Lake Working Group (2009c, page 4) also recommends that standardized OTC derivatives be centrally settled. They would permit customized contracts to remain as bilateral contracts, probably with higher capital charges. See also BCBS (2009a, page 6).} Supervisors should probably charge higher insurance fees or demand more capital against derivatives positions that are settled bilaterally. Removing credit risk exposures from the (vast?) majority of OTC derivatives trades would stabilize the financial system. Regulators and
counterparties could know each firm’s net exposure to various risks, and economic risks could be traded without regard for counterparty performance questions.

As for the case of limiting the scale of bank trading activities, moving OTC settlement to a central counterparty will involve some costs as well as benefits. (For example, see Pirrong (2009) or Duffie and Zhu (2009).) Note that moving a substantial amount of OTC settlement to a central counterparty probably makes that counterparty a systemically important entity (Koeppl and Monnet (2010)). Our financial system already includes similarly systemic facilities, such as the DTCC, the existing futures and options exchanges, and the tri-party repo system. A failure at any of these facilities would have dramatic consequences for financial markets in general.19

Mark Twain allegedly advised people to put all their eggs into one basket, and then to watch the basket very carefully. Our financial infrastructure requires the same sort of tending.

VI. Recommendations

Having considered a variety of regulatory changes aimed at reducing the externalities associated with SIFIs, I am prepared to recommend some specific reforms. Mervyn King (2009, page) points out that

There are only two ways in which the problem can - in logic - be solved. One is to accept that some institutions are “too important to fail” and try to ensure that the probability of those institutions failing, and hence of the need for taxpayer support, is extremely low. The other is to find a way that institutions can fail without imposing unacceptable costs on the rest of society. Any solution must fall into one of those two categories.

No single policy change will adequately address TBTF distortions, nor will any be immune from potential drawbacks. My recommendations aim to address both firm-specific risks and institutional infrastructure. Although the following four innovations offer cost-effective means of reducing the

---

19 The multi-trillion dollar, tri-party repo market involves two private firms (Bank of New York-Mellon and JP Morgan) whose failure would have catastrophic effects for many institutions, particularly since these two clearing banks have incompatible computer systems. In other words, RP transactions could not immediately move from one institution to the other institution in the event of a failure.
distortions created by TBTF financial policies, they cannot guarantee that we will never suffer a repeat of the recent crisis.

1. **Centralized OTC Settlement**

   Encouraging banks to settle standard derivatives contracts (CDS, FX, swaps) on a centralized exchange will eliminate a lot of credit risk from the financial system. These aggregate risk exposures and their possible concentrations cannot presently be observed by supervisors or by market participants. Yet the concentration of OTC derivatives trading among a relatively few SIFIs strongly suggests that many counterparties would be disrupted if one of the main dealers became distressed. There are surely some offsetting costs of clearing against a centralized counterparty, but I believe that the reduction in bilateral credit risk exposures easily justifies a policy of migrating at least standard OTC contracts to a centralized system of margining and settlement.\(^\text{20}\)

2. **Tie Some Supervisory Actions to Market Information**

   Incorporating market-value data into supervisory oversight processes provides a mechanism for disciplining supervisors and limiting their (understandable) inclination to delay taking assertive actions against individual firms. Banks’ incentives become dysfunctional when they are operating with extremely high (market valued) leverage. Book-valued capital ratios inevitably lag changes in market value, particularly when a firm’s condition is deteriorating. Under current accounting standards, book-based solvency measures will “bite” only after market counterparties have lost faith in the troubled firm, by which time regulatory discipline is largely ineffective. By keying regulatory actions (like those in the Prompt Corrective Action regime for U.S. banks) to market capital ratios, supervisory delay can be reduced and bank distress sometimes avoided.

3. **Limit Firms’ Volatile Investments**

\(^{20}\) We should approach this policy with price tools, for example imposing relatively high capital requirements against the credit risk of OTC derivatives. This would leave parties free to negotiate complex or idiosyncratic contracts if the benefits were sufficient to offset the higher capital charges.
SIFIs should be prohibited from engaging in substantial proprietary trading, or from holding large, volatile investments like hedge funds or private equity funds. The appropriate leverage for these investments is generally lower than the appropriate leverage for funding fixed-income claims. Incorporating both sorts of risks within a single institution therefore complicates the problem of computing risk-based capital or insurance fees, as well as identifying an institution’s true risk exposure. At the same time, we must recognize that some of the risk-taking associated with these activities will migrate into large banks’ trading accounts, which have relatively few hard restrictions on their permissible positions. Some risky activities will also migrate to shadow financial institutions.

4. **High Capital Ratios; Contingent Capital**

In pricing bank risk-taking and controlling SIFIs’ externalities, we should place primary reliance on high, suitably risk-based capital standards. Capital is a generic guarantor, good for protecting fixed creditors from any type of loss, and hence for avoiding disruptive bankruptcies. In order to accommodate the measurement errors associated with any fee-based or quantity-based method of controlling SIFIS’ default risks, I propose that risk-based capital requirements include a generous “fudge factor” to afford taxpayers an added level of protection. A separate leverage restriction should protect against the possibility that supervisory risk assessments fail to capture all the relevant risks.

This recommendation will generate three predictable objections. (Maybe more.) First, bankers will assert that higher capital ratios will raise their funding costs and make them unprofitable. Higher loan rates will have detrimental real-sector effects. (See the arguments in Elliott (2010).) Bank lobbyists will further assert that U.S. banks will become internationally uncompetitive if they operate under relatively high capital requirements. Second, corporate governance experts may prefer that managers of opaque asset portfolios be subject to the discipline of “hard” debt claims, rather than “soft” equity claims on the firms’ cash flows (Squam Lake (2009b)). The first argument will carry substantial political weight; the second argument is more debatable, but should not be summarily dismissed. The final, most pragmatic, objection to relying on traditional capital standards to protect the financial system follows
from the recent observations that high book capital does not necessarily assure that market participants will treat the bank as if it is solvent.

How, then, can we introduce substantial risk-bearing capacity into SIFIs, and maintain that capacity, without extraordinarily burdening them?

The concept of contingent capital has received much attention during the past year (BCBS (2009a, page 20), Flannery (2009)). Although it can take various forms, I will discuss the properties of contingent capital in terms of a debt instrument that automatically converts into equity shares if a bank’s capital falls below some critical value. Tax-deductible interest payments will keep a bank’s overall funding cost low during normal times, and provide a hard constraint on managers’ discretion about cash flows. Moreover, shareholders can keep a levered claim on the “upside” outcomes. When losses reduce the bank’s capital, however, the firm is de-levered as debt converts into equity. As losses drive a bank toward its bankruptcy point, contingent capital moves that point further away. This de-levering occurs without raising the specter of an uncertain bankruptcy process.

Many commentators find it plausible that contingent capital could make a valuable contribution to bank capital requirements. However, even the idea’s proponents differ quite substantially in the sort of trigger that should cause conversion. Some have recommended dual triggers: conversion would occur only if the issuer and the financial system were under pressure at the same time (MacDonald (2010), Squam Lake (2009a)). The problem I see with the systemic trigger is that supervisors are unlikely to permit even one “systemically important” firm to fail. If any SIFI’s imminent failure would elicit government support, then it requires a mechanism for de-levering outside of bankruptcy.

Another important difference among analysts has been the nature of the triggering event. Some would use a book capital ratio (for example, “the ratio of Tier 1 capital to risk-adjusted assets” suggested by the Squam Lake group (2009a)), while others would require supervisors to trigger the conversion based on their contemporaneous information. As shown above, a Tier 1 trigger is unlikely to be tripped

---

21 Contingent capital could take other forms, including convertible preferred shares, forgivable (“CAT”) bonds, insurance payments, or pre-arranged equity put options.
even under extreme conditions.\footnote{Lloyd’s 2009 issue of “Enhanced Credit Notes” specifies a share conversion price above the likely market price associated with its book-value trigger. A high conversion price imposes losses on CC at conversion.\footnote{Shareholders enjoy a corresponding gain.} These bond investors will hold substantial equity risk as the issuer’s capital value approaches the trigger. Bondholders short the issuer’s stock to hedge this risk, but the need for this talent narrows the contingent capital bonds’ market. Building equity risk into CC bonds would probably discourage traditional bond investors from holding contingent capital bonds, making it more difficult to insert a large volume of contingent capital into banks. And if the CC tranche of bank financing is too small, its conversion will do little to forestall a run or a default.} Supervisors would be reluctant to force a conversion, for fear of signaling negative private information to market participants. Furthermore, supervisors already have the power to force a troubled firm to issue new equity, but have rarely taken that step. Only a market-valued equity trigger (e.g. “equity value below x% of the most recent quarter-end’s\footnote{Ideally, the denominator would also be valued at market prices, but daily asset values are not readily observable.} risk-weighted assets”) can elicit de-levering as soon as the market becomes concerned about the firm’s solvency.

A bond that “converts” into equity when the firm is experiencing the first signs of financial distress sounds as if bondholders should suffer large, share-value-related losses. This potential for default losses would probably limit the number of interested investors. However, the contingent bondholders’ loss at conversion depends on the share price at which the bond converts. In the limit, if bonds convert at the shares’ contemporaneous market value, bondholders suffer no loss at conversion. The bonds are quite safe, and can be sold into a broad market of traditional bond investors. Moreover, supervisors should prefer a relatively high trigger value, so that converted bonds augment the capital of a firm with positive going concern value. A higher trigger reduces the probability that contingent debt will suffer a credit loss.

\footnote{A tangible common equity (TCE) ratio might be worth considering, if the evidence shows that it reflects asset value changes more accurately than Tier I capital measures do. Even so, TCE’s tendency to lag market value measures means that a TCE trigger should be higher than it would need to be for a market value trigger.} A bond that “converts” into equity when the firm is experiencing the first signs of financial distress sounds as if bondholders should suffer large, share-value-related losses. This potential for default losses would probably limit the number of interested investors. However, the contingent bondholders’ loss at conversion depends on the share price at which the bond converts. In the limit, if bonds convert at the shares’ contemporaneous market value, bondholders suffer no loss at conversion. The bonds are quite safe, and can be sold into a broad market of traditional bond investors. Moreover, supervisors should prefer a relatively high trigger value, so that converted bonds augment the capital of a firm with positive going concern value. A higher trigger reduces the probability that contingent debt will suffer a credit loss.

\footnote{Rabobank’s 2010 convertible debt issue also imposes a loss on shareholders if the book-capital trigger is tripped: the bonds are immediately repaid at a 75\% discount.} A bond that “converts” into equity when the firm is experiencing the first signs of financial distress sounds as if bondholders should suffer large, share-value-related losses. This potential for default losses would probably limit the number of interested investors. However, the contingent bondholders’ loss at conversion depends on the share price at which the bond converts. In the limit, if bonds convert at the shares’ contemporaneous market value, bondholders suffer no loss at conversion. The bonds are quite safe, and can be sold into a broad market of traditional bond investors. Moreover, supervisors should prefer a relatively high trigger value, so that converted bonds augment the capital of a firm with positive going concern value. A higher trigger reduces the probability that contingent debt will suffer a credit loss.
Many regulators, from many countries, have recently asserted that banks must hold more capital going forward. This will tend to reduce bank profits, *ceteris paribus*. But holding some of the increased capital in the form of contingent capital bonds provides a win-win outcome. If banks can hold some capital in the form of contingently convertible, tax-deductible debt, their cost of funds won’t rise as much as if common equity provided all of their downside risk absorption. Regulators and taxpayers can then get more downside risk absorption in the ensuing bargaining game than they would get if all capital was compensated with pre-tax dollars.

V. Conclusion

In the wake of TBTF-related losses associated with the recent financial crisis, many commentators have proposed supervisor changes designed to mitigate the effects of large, failure-proof financial institutions on our economy’s financial and real sectors. If the TBTF problem has an easy, straightforward solution, I’ve yet to identify it. This paper has reviewed those proposals and recommended four that seems worth pursuing. Of course these institutional and regulatory changes won’t eliminate TBTF distortions. Nor can reforms assure that the financial system will never suffer another catastrophic event. However, these recommended changes will reduce the probability of distortions of TBTF operations and reduce the impact of a SIFI’s failure.

As a final comment, it is important to emphasize that my recommended changes will tend to drive some risky activities outside the banking system. Supervisors (perhaps the systemic risk regulator) must therefore remain alert for the formation of non-bank SIFIs, and incorporate them into their supervisory oversight. A difficult process. As usual, prudential regulation will need to evolve with the institutions to which it applies. This is probably the most daunting challenge supervisors face, going forward.
References


Bernanke, Ben S., “Preserving a Central Role for Community Banking” (March 20, 2010).


________ (2009), “Stabilizing Large Financial Institutions with Contingent Capital Certificates”, University of Florida working paper.


Johnson, Simon (2009), “Above all, we need to encourage or, most likely, force the large insolvent banks to break up”, (Feb. 23, 2009). (last accessed on April 12, 2010).

King, Mervyn, Edinburgh Speech (October 20, 2009).
Thorsten V. Koeppl, Thorsten V. and Cyril Monnet (2010), “The Emergence and Future of Central Counterparties”.

Kuritzkes, Andrew and Hal Scott, “Markets are the best judge of bank capital,” Financial Times (September 23, 2009).


O’Driscoll, Gerald (2009), “If a bank is too big to fail, then it is simply too big”, Wall Street Journal (Feb. 23, 2009).


Pierce, James, The Future of Banking (1991), Yale University Press.


Figure 1:

Figure 2:

Figure 3:
Figure 4: Book Capital Ratios vs. Market Solvency Indicators

Notes: Market value and book value ratios are simple means for 18 FIs that participated in the SCAP, excluding GMAC. CDS spreads are simple means of available data.
Figure 5: Market vs. Book Values for Specific Banks

Citigroup

BankAmerica

Tier1/RWA
MVEQ/BVA
Wachovia

JPMorgan/Chase